

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 26, 2002, 20:33:06 ; Search time 1798 Seconds
(without alignments)
8065.678 Million cell updates/sec

Title: US-09-840-795-18_COPY_78_770

Perfect score: 693
Sequence: 1 atgagttgcgaagaataatga.....agcagcaggggcctgaatg 693

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

Genemdb1:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
------------	-------------	-------	--------	----	-------------

1	611.4	88.2	1162	6	AX127536	AX127536 Sequence
2	609.4	87.9	894	6	AF298812	AF298812 Homo sapi
3	446.6	64.4	891	6	AX127563	AX127563 Sequence
4	446.4	64.4	807	6	AX127538	AX127538 Sequence
5	413.2	59.6	1081	6	AX127585	AX127585 Sequence
6	393.2	56.7	534	6	AX127556	AX127556 Sequence
7	393.2	56.7	1200	6	AX127557	AX127557 Sequence
8	317.4	45.8	801	6	AX127565	AX127565 Sequence
9	315.4	45.5	519	6	AX127574	AX127574 Sequence
10	257	37.1	528	6	AX127568	AX127568 Sequence
11	220.4	31.8	529	6	AX127566	AX127566 Sequence
12	181.4	26.2	192505	6	AL353136	AL353136 Human DNA
13	181.4	26.2	205736	2	AC023560	AC023560 Homo sapi
14	163.2	23.5	744	10	AB040433	AB040433 Mus muscu
15	163.2	23.5	886	10	AF167553	AF167553 Mus muscu
16	163.2	23.5	1075	10	AF173166	AF173166 Mus muscu
17	163.2	23.5	1678	10	AF167552	AF167552 Mus muscu
18	163.2	23.5	1914	6	AX213199	AX213199 Sequence
19	163.2	23.5	1914	10	AF247000	AF247000 Mus muscu
20	163.2	23.5	3964	10	AB040432	AB040432 Mus muscu
21	156.4	22.6	1489	9	AF167555	AF167555 Homo sapi
22	154.8	22.3	1254	6	AX151748	AX151748 Sequence
23	154.8	22.3	1325	6	AX213197	AX213197 Sequence
24	154.8	22.3	1325	9	AF246999	AF246999 Homo sapi
25	154.8	22.3	1364	9	AB040434	AB040434 Homo sapi
26	154.8	22.3	1660	6	AX213195	AX213195 Sequence
27	154.8	22.3	1660	9	AF246998	AF246998 Homo sapi
28	151.2	21.8	893	6	AR171888	AR171888 Sequence
29	150.8	21.8	205736	2	AC023560	AC023560 Homo sapi
30	144.2	20.8	43294	2	AC100492	AC100492 Mus muscu
31	132.8	19.2	591	10	AF167554	AF167554 Mus muscu
32	132.8	19.2	623	6	AR171889	AR171889 Sequence
33	120	17.3	190747	2	AC096891	AC096891 Rattus no
34	96	13.9	431	6	AX127571	AX127571 Sequence
35	95.2	13.7	401	6	AX127567	AX127567 Sequence
36	77.4	11.2	190747	2	AC096891	AC096891 Rattus no
37	76	11.0	43294	2	AC100492	AC100492 Mus muscu
38	75.8	10.9	143420	9	AL161422	AL161422 Human DNA
39	51.2	7.4	67259	2	AC100501	AC100501 Mus muscu
40	44.6	6.4	125020	9	AF429315	AF429315 Homo sapi
41	40.2	5.8	1260	10	AF411030	AF411030 Mus muscu
42	40.2	5.8	60709	2	AC107923	AC107923 Homo sapi
43	40.2	5.8	60709	2	AC107923	AC107923 Homo sapi
44	39	5.6	123551	2	AC005809	AC005809 Homo sapi
45	39	5.6	125020	9	AF429315	AF429315 Homo sapi

ALIGNMENTS

RESULT 1
AX127536 1162 bp DNA linear PAT 15-MAY-2001
LOCUS AX127536
DEFINITION Sequence 1 from Patent WO0130850.
ACCESSION AX127536
VERSION AX127536.1 GI:14134241
KEYWORDS
SOURCE
ORGANISM human.
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1162)
AUTHORS Xu,W., Lofton-Day,C.E., Henne,R., Yao,Y., Novak,J.E., Foster,D.C.
and Yee,D.P.
TITLE Umlr polypeptides
JOURNAL Patent: WO 0130850-A 1 03-MAY-2001;
ZymoGenetics, Inc. (US)
FEATURES
source location/Qualifiers
1..1162
/organism="Homo sapiens"
/db_xref="taxon:9606"
104..913
/note="unnamed protein product"

/codon_start=1
/protein_id="CAC38899.1"
/db_xref="GI:14134242"
/translation="MDCQENYWDQRCVTCORCGPGLSKDCGYEGEGDAYCTAC
PPRRKSSMHHNCOSCTICAVINRVKVCATSNVACDCLPREFRKIRIGLQDO
ECIPCTKOTPTSEVQCAFOLSLVADAPVPOPATLVALLVETLAFGLFEL
YCOFENRHCORGLLOFADKTKAKESLPPVPSKETSSESVSEFTMACTSESHSH
WVHSPICECTLDLQKFFSSASATGAEFLGNTVSTEDRLINVPFVPSP"
BASE COUNT 255 a 327 c 314 g 266 t
ORIGIN

Query Match 88.2%; Score 611.4; DB 6; Length 1162;
Best Local Similarity 99.0%; Pred. No. 4.4e-167;
Matches 615; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 ATGATTTGCCAAGAAATGACTGTGGACCAATGGGAGCGTGTCTACCTGCCAACGG 60
|||||
DB 104 ATGATTTGCCAAGAAATGACTGTGGACCAATGGGAGCGTGTCTACCTGCCAACGG 163
|||||
OY 61 TGTGTCTCTGGACAGAGGTATCCAAAGATTGTGTTATGGAGAGGTGGAGATGCTTAC 120
|||||
DB 164 TGTGTCTCTGGACAGAGGTATCCAAAGATTGTGTTATGGAGAGGTGGAGATGCTTAC 223
|||||
OY 121 TGCACAGCGCTGCTCTCCGTCAGTACAAAGACAGCTGGGGCCACACAAATGTCAGAGT 180
|||||
DB 224 TGCACAGCGCTGCTCTCCGTCAGTACAAAGACAGCTGGGGCCACACAAATGTCAGAGT 283
|||||
OY 181 TGCATCACCTGTGCTGTCAATCAATGCTGTTCAGAAAGTCACTGCACAGCTTCTTAAT 240
|||||
DB 284 TGCATCACCTGTGCTGTCAATCAATGCTGTTCAGAAAGTCACTGCACAGCTTCTTAAT 343
|||||
OY 241 GCTGTCTGTGGGGAGCTGTTTGCCCAAGTCTTACCAAGACACGATTGGAGCGCCGACG 300
|||||
DB 344 GCTGTCTGTGGGGAGCTGTTTGCCCAAGTCTTACCAAGACACGATTGGAGCGCCGACG 403
|||||
OY 301 GACCAAGATGATCCCGTCGACAGACAGACCCCACTCTGAGATTCAATGTGCTTTC 360
|||||
DB 404 GACCAAGATGATCCCGTCGACAGACAGACCCCACTCTGAGATTCAATGTGCTTTC 463
|||||
OY 361 CAGTTGAGCTTAGTGGAGGCAAGTGCACCAAGTGCCTTCAGAGGCGCACACTTGT 420
|||||
DB 464 CAGTTGAGCTTAGTGGAGGCAAGTGCACCAAGTGCCTTCAGAGGCGCACACTTGT 523
|||||
OY 421 GCACGTGTAGAGCGCTGTAGTGGTGTACCTTGCGCTTCCTGGGGCTCTTCTCTC 480
|||||
DB 524 GCACGTGTAGAGCGCTGTAGTGGTGTACCTTGCGCTTCCTGGGGCTCTTCTCTC 583
|||||
OY 481 TACTGCACAGCTTCTTCAACAGACATTCACAGCGTGCAGAGTTGCTGACAGTTGAGGCT 540
|||||
DB 584 TACTGCACAGCTTCTTCAACAGACATTCACAGCGTGCAGAGTTGCTGACAGTTGAGGCT 643
|||||
OY 541 GATTAACAGCAAGAGAGATTCCTTCCCGTGCACCAAGAGAGACAGCACTTGT 600
|||||
DB 644 GATTAACAGCAAGAGAGATTCCTTCCCGTGCACCAAGAGAGAGACAGCACTTGT 703
|||||
OY 601 GAGTCCCAAGTCTCTTGGGCG 621
|||||
DB 704 GAGTCCCAAGAGTCTTTTAC 724
|||||

RESULT 2 894 bp mRNA linear PRI 01-NOV-2000
AF298812
LOCUS Homo sapiens X-linked ectodysplasin-A2 receptor (XEDAR) mRNA,
DEFINITION complete cds.
ACCESSION AF298812
VERSION AF298812.1 GI:11066914
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 894)

AUTHORS Van, M., Wang, L.C., Hymowitz, S.G., Schillbach, S., Lee, J., Goddard, A.,
de Vos, A.M., Gao, W.-Q. and Dixit, V.M.
TITLE Two-amino acid molecular switch in an epithelial morphogen that
regulates binding to two distinct receptors
JOURNAL Science 290 (5491), 523-527 (2000)
MEDLINE 20495245
PUBMED 11039935
REFERENCE 2 (bases 1 to 894)
AUTHORS Van, M., Schillbach, S., Goddard, A. and Dixit, V.M.
TITLE Direct Submission
JOURNAL Submitted (23-AUG-2000) Molecular Oncology, Genentech Inc., 1 DNA
Way, South San Francisco, CA 94080, USA
Location/Qualifiers

FEATURES

source
1..894
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
1..894
/gene="XEDAR"
1..894
/gene="XEDAR"
/note="member of the TNFR superfamily"
/codon_start=1
/product="X-linked ectodysplasin-A2 receptor"
/protein_id="AAC28761.1"
/db_xref="GI:11066915"
/translation="MDCQENYWDQRCVTCORCGPGLSKDCGYEGEGDAYCTAC
PPRRKSSMHHNCOSCTICAVINRVKVCATSNVACDCLPREFRKIRIGLQDO
ECIPCTKOTPTSEVQCAFOLSLVADAPVPOPATLVALLVETLAFGLFEL
YCOFENRHCORGLLOFADKTKAKESLPPVPSKETSSESVSEFTMACTSESHSH
WVHSPICECTLDLQKFFSSASATGAEFLGNTVSTEDRLINVPFVPSP"
NVESTEDRLINVPFVPSP"

BASE COUNT 207 a 256 c 238 g 193 t
ORIGIN

Query Match 87.9%; Score 609.4; DB 9; Length 894;
Best Local Similarity 99.8%; Pred. No. 1.7e-166;
Matches 610; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ATGATTTGCCAAGAAATGACTGTGGACCAATGGGAGCGTGTCTACCTGCCAACGG 60
|||||
DB 1 ATGATTTGCCAAGAAATGACTGTGGACCAATGGGAGCGTGTCTACCTGCCAACGG 60
|||||
OY 61 TGTGTCTCTGGACAGAGGTATCCAAAGATTGTGTTATGGAGAGGTGGAGATGCTTAC 120
|||||
DB 61 TGTGTCTCTGGACAGAGGTATCCAAAGATTGTGTTATGGAGAGGTGGAGATGCTTAC 120
|||||
OY 121 TGCACAGCGCTGCTCTCCGTCAGTACAAAGACAGCTGGGGCCACACAAATGTCAGAGT 180
|||||
DB 121 TGCACAGCGCTGCTCTCCGTCAGTACAAAGACAGCTGGGGCCACACAGATGTCAGAGT 180
|||||
OY 181 TGCATCACCTGTGCTGTCAATCAATGCTGTTCAGAAAGTCACTGCACAGCTTCTTAAT 240
|||||
DB 181 TGCATCACCTGTGCTGTCAATCAATGCTGTTCAGAAAGTCACTGCACAGCTTCTTAAT 240
|||||
OY 241 GCTGTCTGTGGGGAGCTGTTTGCCCAAGTCTTACCAAGACACGATTGGAGCGCCGACG 300
|||||
DB 241 GCTGTCTGTGGGGAGCTGTTTGCCCAAGTCTTACCAAGACACGATTGGAGCGCCGACG 300
|||||
OY 301 GACCAAGATGATCCCGTCGACAGACAGACCCCACTCTGAGATTCAATGTGCTTTC 360
|||||
DB 301 GACCAAGATGATCCCGTCGACAGACAGACCCCACTCTGAGATTCAATGTGCTTTC 360
|||||
OY 361 CAGTTGAGCTTAGTGGAGGCAAGTGCACCAAGTGCCTTCAGAGGCGCACACTTGT 420
|||||
DB 361 CAGTTGAGCTTAGTGGAGGCAAGTGCACCAAGTGCCTTCAGAGGCGCACACTTGT 420
|||||
OY 421 GCACGTGTAGAGCGCTGTAGTGGTGTACCTTGCGCTTCCTGGGGCTCTTCTCTC 480
|||||
DB 421 GCACGTGTAGAGCGCTGTAGTGGTGTACCTTGCGCTTCCTGGGGCTCTTCTCTC 480
|||||
OY 481 TACTGCACAGCTTCTTCAACAGACATTCACAGCGTGCAGAGTTGCTGACAGTTGAGGCT 540
|||||

Dh 481 TACTGACAGCAGTTCTTCAACAGACATTGCGAGGTGGAGTTTGTCTGCACTTTGAGGCT 540
Qy 541 GATTAACAGCAAGAGAGGAAATCTCTTCCCGTGGCCACCAGCAAGAGAGCAGTGGCT 600
Dh 541 GATTAACAGCAAGAGAGGAAATCTCTTCCCGTGGCCACCAGCAAGAGAGCAGTGGCT 600
Qy 601 GAGTCCCAAGT 611
Dh 601 GAGTCCCAAGT 611

RESULT 3
AX127563

LOCUS AX127563 891 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 28 from Patent WO0130850.
ACCESSION AX127563
VERSION AX127563.1 GI:14134260
KEYWORDS

SOURCE synthetic construct.

ORGANISM synthetic construct.

REFERENCE 1 (bases 1 to 891)
AUTHORS Xu, W., Lofton-Day, C.E., Henne, R., Yao, Y., Novak, J.E., Foster, D.C.TITLE Umlr polypeptides
JOURNAL Patent: WO 0130850-A 28 03-MAY-2001;
ZymoGenetics, Inc. (US)FEATURES
source location/Qualifiers1..891 /organism="synthetic construct"
/db_xref="taxon:32630"misc_feature /note="degenerate sequence"
1..891BASE COUNT 141 a 105 c 150 g 112 t 383 others
ORIGINQuery Match 64.4%; Score 446.6; DB 6; Length 891;
Best Local Similarity 58.6%; Pred. No. 4,5e-119;

Matches 360; Conservative 147; Mismatches 107; Indels 0; Gaps 0;

Qy 481 TACTGACAGCAGTTCTTCAACAGACATTGCGAGGTGGAGTTTGTCTGCACTTTGAGGCT 540
Dh 481 TATTCYAAARCAKRTTYYTAAATGCAATGTCARNGNGNGNNTYNTNCATTTGARGCN 540
Qy 541 GATTAACAGCAAGAGAGGAAATCTCTTCCCGTGGCCACCAGCAAGAGAGCAGTGGCT 600
Dh 541 GATTAACAGCAAGAGAGGAAATCTCTTCCCGTGGCCACCAGCAAGAGAGCAGTGGCT 600
Qy 601 GAGTCCCAAGTCTC 614
Dh 601 GAGTCCCAAGTCTC 614

RESULT 4
AX127538

LOCUS AX127538 807 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 3 from Patent WO0130850.
ACCESSION AX127538
VERSION AX127538.1 GI:14134243
KEYWORDS

SOURCE synthetic construct.

ORGANISM synthetic construct.

REFERENCE 1 (bases 1 to 807)
AUTHORS Xu, W., Lofton-Day, C.E., Henne, R., Yao, Y., Novak, J.E., Foster, D.C.TITLE Umlr polypeptides
JOURNAL Patent: WO 0130850-A 3 03-MAY-2001;
ZymoGenetics, Inc. (US)FEATURES
source location/Qualifiers1..807 /organism="synthetic construct"
/db_xref="taxon:32630"misc_feature /note="degenerate sequence"
1..807BASE COUNT 125 a 93 c 142 g 102 t 345 others
ORIGINQuery Match 64.4%; Score 446.4; DB 6; Length 807;
Best Local Similarity 55.4%; Pred. No. 5.1e-119;

Matches 370; Conservative 156; Mismatches 142; Indels 0; Gaps 0;

Qy 1 ATGAGTTGCCAAGAAATGAGTACTGGACCAATGGGAGCGGTGTGTCACCTGCCAAGCG 60
Dh 1 ATGAGTTGCCAAGAAATGAGTACTGGACCAATGGGAGCGGTGTGTCACCTGCCAAGCG 60
Qy 61 TGTGTCCTGGACAGAGGCTATCCAGAGTGTGTTATGAGAGGTGGAGTGCCTAC 120
Dh 61 TGTGTCCTGGACAGAGGCTATCCAGAGTGTGTTATGAGAGGTGGAGTGCCTAC 120
Qy 121 TGCACAGCCTGCCCTCTCGCAGGTACCAAGAGCGTGGGCGCACCAATGTCACAGT 180
Dh 121 TGCACAGCCTGCCCTCTCGCAGGTACCAAGAGCGTGGGCGCACCAATGTCACAGT 180
Qy 181 TGCATCAGCTGTCTGTCAATCGTGTTCAGAGGTCAACTGCACAGCTACTCTTAAT 240
Dh 181 TGCATCAGCTGTCTGTCAATCGTGTTCAGAGGTCAACTGCACAGCTACTCTTAAT 240
Qy 241 GCTGTCTGGGAGCTGTGGCCAGGTCTACCGAAAGCAGCATTTGAGGCTTCAG 300
Dh 241 GCTGTCTGGGAGCTGTGGCCAGGTCTACCGAAAGCAGCATTTGAGGCTTCAG 300
Qy 301 GACCAAGAGTGCATCCCGTCAGCAAGAGCAGCCCACTGTGAGTTCAATGGCTTC 360
Dh 301 GACCAAGAGTGCATCCCGTCAGCAAGAGCAGCCCACTGTGAGTTCAATGGCTTC 360
Qy 361 CACTTGAAGCTTAATGAGAGGAGATGCAACCAAGTGCCTGCCCTCAGAGGCGCACTTGT 420
Dh 361 CACTTGAAGCTTAATGAGAGGAGATGCAACCAAGTGCCTGCCCTCAGAGGCGCACTTGT 420
Qy 421 GCACGTGTAGAGCGCTGCTAGTGTGTACCTGCGCTTCTGCGGCTCTTCTCTC 480
Dh 421 GCACGTGTAGAGCGCTGCTAGTGTGTACCTGCGCTTCTGCGGCTCTTCTCTC 480

Qy 481 TACTGACAGCAGTTCTTCAACAGACATTGCGAGGTGGAGTTTGTCTGCACTTTGAGGCT 540
Dh 481 TATTCYAAARCAKRTTYYTAAATGCAATGTCARNGNGNGNNTYNTNCATTTGARGCN 540
Qy 541 GATTAACAGCAAGAGAGGAAATCTCTTCCCGTGGCCACCAGCAAGAGAGCAGTGGCT 600
Dh 541 GATTAACAGCAAGAGAGGAAATCTCTTCCCGTGGCCACCAGCAAGAGAGCAGTGGCT 600
Qy 601 GAGTCCCAAGTCTC 614
Dh 601 GAGTCCCAAGTCTC 614

Db 421 GCNTYNTGMSWMSNYTNTNGTNGTNTTYACNTYNGCNTTYTYTNGNNTNTTYTYTN 480

QY 481 TACTGCAAGCAGTCTTCAACAGACATGCGCGCTGGAGGTTTCTGCTGAGTTTACGCT 540

Db 481 TAYTYAARCAATTTTAAATGNCAYTGCAKRMGNGGNGNTYNTTCARTTYAGRGCN 540

QY 541 GATTAACAGCAAGAGAGAAATCTCTCTCCCGCTGCGCACCCAGCAGAGAGACAGTCT 600

Db 541 GAYAAACACGCAAGAGAGAAATCTCTCTCCCGCTGCGCACCCAGCAGAGAGACAGTCT 600

QY 601 GAGTCCCAAGTCTCTGGGCCCCCTGGCAGCCCTTGGCCAGTTGTTCTCTGAGACTCTGTT 660

Db 601 GARMSNCARGAARMSNTTYACNATGCGNMSNTGYACNMSNGARMSNCAYMSNCAYGGTN 660

QY 661 CCTATACC 668

Db 661 CAYMSNC 668

RESULT 5

AX127585 LOCUS AX127585 1081 bp DNA linear PAT 15-MAY-2001

DEFINITION Sequence 50 from Patent W00130850.

ACCESSION AX127585

VERSION AX127585.1 GI:14134277

KEYWORDS

SOURCE synthetic construct.

ORGANISM synthetic construct.

REFERENCE 1 (bases 1 to 1081)

AUTHORS Xu,W., Lofton-Day,C.E., Henne,R., Yao,Y., Novak,J.E., Foster,D.C. and Yee,D.P.

TITLE Umlr polypeptides

JOURNAL Patent: WO 0130850-A 50 03-MAY-2001;

FEATURES

source Location/Qualifiers

1..1081

/organism="synthetic construct"

/db_xref="taxon:32630"

/note="DNA construct"

BASE COUNT 248 a 311 c 306 g 216 t

ORIGIN

Query Match 59.6%; Score 413.2; DB 6; Length 1081;

Best Local Similarity 99.3%; Pred. No. 2.5e-109;

Matches 415; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGATTTGCCAAGAAATAGTACTGAGCAACAATGGGAGCGGTGTCTACCTGCCAACGG 60

Db 1 ATGATTTGCCAAGAAATAGTACTGAGCAACAATGGGAGCGGTGTCTACCTGCCAACGG 60

QY 61 TGTGTGCTCGAGACAGAGAGTATCCAGAGATTTGTTATGAGAGAGGTGAGATGCTTAC 120

Db 61 TGTGTGCTCGAGACAGAGAGTATCCAGAGATTTGTTATGAGAGAGGTGAGATGCTTAC 120

QY 121 TGCACAGCTGCGCTCTCTCGCAGGTACAAAGAGAGCTGGGGCCACCAAAATGTGAGAT 180

Db 121 TGCACAGCTGCGCTCTCTCGCAGGTACAAAGAGAGCTGGGGCCACCAAAATGTGAGAT 180

QY 181 TGCATCACCCTGTGCTGTCAATGCTGTTACAGAGCTCACTGACAGCTTACTTAAT 240

Db 181 TGCATCACCCTGTGCTGTCAATGCTGTTACAGAGCTCACTGACAGCTTACTTAAT 240

QY 241 GCTGTGTGGGGAGACTGTTTGGCCAGGTCTACCGAAACACGATTTGAGAGGCGTGCAG 300

Db 241 GCTGTGTGGGGAGACTGTTTGGCCAGGTCTACCGAAACACGATTTGAGAGGCGTGCAG 300

QY 301 GACCAAGAGTATCCGCTGCAAGAGACAGACCCCACTCTGAGTTCAATGTGCTTC 360

Db 301 GACCAAGAGTATCCGCTGCAAGAGACAGACCCCACTCTGAGTTCAATGTGCTTC 360

QY 361 CAGTTGAGCTTAGTGAGGACAGATGACCCACAGTCCCCCTGAGAGGCCACACATTG 418

Db 361 CAGTTGAGCTTAGTGAGGACAGATGACCCACAGTCCCCCTGAGAGGTGACCGTTG 418

RESULT 6

AX127556 LOCUS AX127556 534 bp DNA linear PAT 15-MAY-2001

DEFINITION Sequence 21 from Patent W00130850.

ACCESSION AX127556

VERSION AX127556.1 GI:14134254

KEYWORDS

SOURCE synthetic construct.

ORGANISM synthetic construct.

REFERENCE 1 (bases 1 to 534)

AUTHORS Xu,W., Lofton-Day,C.E., Henne,R., Yao,Y., Novak,J.E., Foster,D.C. and Yee,D.P.

TITLE Umlr polypeptides

JOURNAL Patent: WO 0130850-A 21 03-MAY-2001;

FEATURES

source Location/Qualifiers

1..534

/organism="synthetic construct"

/db_xref="taxon:32630"

BASE COUNT 125 a 136 c 155 g 118 t

ORIGIN

Query Match 56.7%; Score 393.2; DB 6; Length 534;

Best Local Similarity 99.2%; Pred. No. 1.6e-103;

Matches 395; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 GATTGCCAAGAAATAGTACTGGAGCAATGGGAGCGGTGTCTACCTGCCAACGGTGT 63

Db 109 GATTGCCAAGAAATAGTACTGGAGCAATGGGAGCGGTGTCTACCTGCCAACGGTGT 168

QY 64 GGTCTGAGACAGAGACTTCCAGAGATTTGTTATGAGAGGGGTGAGATGCTTACG 123

Db 169 GGTCTGAGACAGAGACTTCCAGAGATTTGTTATGAGAGGGGTGAGATGCTTACG 228

QY 124 ACAGCCTGCGCTCTCGCAGAGTACAAAGACAGCTGGGCGCACACAATGTGAGATTGC 183

Db 229 ACAGCCTGCGCTCTCGCAGAGTACAAAGACAGCTGGGCGCACACAATGTGAGATTGC 288

QY 184 ATCACCCTGTGCTGTCAATGCTTTCAGAGATTCACAGTACAGTACCTTAATGCT 243

Db 289 ATCACCCTGTGCTGTCAATGCTTTCAGAGATTCACAGTACAGTACCTTAATGCT 348

QY 244 GCTGTGGGAGACTGTTTGGCCAGGTCTTACCGAAAGACACGATTTGAGAGGCGCTGAGGNC 303

Db 349 GCTGTGGGAGACTGTTTGGCCAGGTCTTACCGAAAGACACGATTTGAGAGGCGCTGAGGNC 408

QY 304 CAAGAGTCAATCCGCTGCAAGAGACAGACCCCACTCTGAGTTCAATGTGCTTCAG 363

Db 409 CAAGAGTCAATCCGCTGCAAGAGACAGACCCCACTCTGAGTTCAATGTGCTTCAG 468

QY 364 TTGAGCTTAGTGAGGACAGATGACCCACAGTGGCCCC 401

Db 469 TTGAGCTTAGTGAGGACAGATGACCCACAGTGGAGCC 506

RESULT 7

AX127557 LOCUS AX127557 1200 bp DNA linear PAT 15-MAY-2001

DEFINITION Sequence 22 from Patent W00130850.

ACCESSION AX127557

VERSION AX127557.1 GI:14134255

KEYWORDS

SOURCE synthetic construct.

ORGANISM synthetic construct.

REFERENCE 1 (bases 1 to 1200)

AUTHORS Xu,W., Lofton-Day,C.E., Henne,R., Yao,Y., Novak,J.E., Foster,D.C. and Yee,D.P.

TITLE Umlr polypeptides

JOURNAL	Patent: WO 0130850-A 22 03-MAY-2001;									
FEATURES	ZymoGenetics, Inc. (US)									
SOURCE	Location/Qualifiers 1. .1200 /organism="synthetic construct" /db_xref="taxon:32630" /note="construct"									
BASE COUNT	290 a	358 c	330 g	222 t						
ORIGIN										
Query Match	56.7%;	Score 393.2;	DB 6;	Length 1200;						
Best Local Similarity	99.2%;	Pred. No. 1.7e-103;								
Matches 395;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;						
OY	4	GATTGCCAAGAAATGAGTACTGCGACCAATGCGACGGTGTCTACCTGCCACGGTGT	63							
Db	109	GATTGCCAAGAAATGAGTACTGCGACCAATGCGACGGTGTCTACCTGCCACGGTGT	168							
OY	64	GGTCTTGACAGAGAGCTATGCCAAGATTGTGGTATGAGAGCGGAGATCCCTACTGC	123							
Db	169	GGTCTTGACAGAGAGCTATGCCAAGATTGTGGTATGAGAGCGGAGATCCCTACTGC	228							
OY	124	ACAGCCTCCCTCTCTCGAGAGTACAAAAGACAGCTGGGGCCACACAAATGTGAGATTGC	183							
Db	229	ACAGCCTCCCTCTCTCGAGAGTACAAAAGACAGCTGGGGCCACACAAATGTGAGATTGC	288							
OY	184	ATCACCTGTGCTGTCATCAATGCTGTTCAGAGGTGCACACTGCACAGCTACCTTAATGCT	243							
Db	289	ATCACCTGTGCTGTCATCAATGCTGTTCAGAGGTGCACACTGCACAGCTACCTTAATGCT	348							
OY	244	GTCGTGTGGGACCTGTTTCCCGAGCTTCTACCGAAAGACACGCATTGAGAGCCTCGAGAGAC	303							
Db	349	GTCGTGTGGGACCTGTTTCCCGAGCTTCTACCGAAAGACACGCATTGAGAGCCTCGAGAGAC	408							
OY	304	CAAGAGTCAATCCGTCGACAGAGACAGACCCACCTCTGAGGTTCAATGTGCTTTCAG	363							
Db	409	CAAGAGTCAATCCGTCGACAGAGAGACCCACCTCTGAGGTTCAATGTGCTTTCAG	468							
OY	364	TTGAGCTTAGTGGAGGCGAGATGCACCCAGTGGCCCCC	401							
Db	469	TTGAGCTTAGTGGAGGCGAGATGCACCCAGTGGAGCC	506							
RESULT 8										
LOCUS	AX127565	801 bp	DNA	Linear	PAT 15-MAY-2001					
DEFINITION	Sequence 30 from Patent W00130850.									
ACCESSION	AX127565									
VERSION	AX127565.1	GI:14134261								
KEYWORDS										
SOURCE	synthetic construct.									
ORGANISM	synthetic construct.									
REFERENCE	1 (bases 1 to 801)									
AUTHORS	Xu,W., Lofton-Day,C.E., Henne,R., Yao,Y., Novak,J.E., Foster,D.C. and Yee,D.P.									
TITLE	Umlr polypeptides									
JOURNAL	Patent: WO 0130850-A 30 03-MAY-2001;									
FEATURES	ZymoGenetics, Inc. (US)									
SOURCE	Location/Qualifiers 1. .801 /organism="synthetic construct" /db_xref="taxon:32630" /note="degenerate sequence"									
misc_feature	1. .801									
BASE COUNT	132 a	100 c	140 g	88 t	341 others					
ORIGIN										
Query Match	45.8%;	Score 317.4;	DB 6;	Length 801;						
Best Local Similarity	59.6%;	Pred. No. 2e-81;								
Matches 265;	Conservative 97;	Mismatches 83;	Indels 0;	Gaps 0;						

OY	1	ATGGATGTCCCAAGAAATGAGTACATGGGACCATAAGGGAGCGGTGTGTACTCCTGCCAACGG	60
Db	1	ATGGATGTCCCAAGAAATGAGTACATGGGACCATAAGGGAGCGGTGTGTACTCCTGCCAACGG	60
OY	1	ATGGATGTCCCAAGAAATGAGTACATGGGACCATAAGGGAGCGGTGTGTACTCCTGCCAACGG	60
Db	1	ATGGATGTCCCAAGAAATGAGTACATGGGACCATAAGGGAGCGGTGTGTACTCCTGCCAACGG	60
OY	61	TGTGTCTCTGCACAGACCTCATCCAAGATTGTGGTTATGACAGAGGTGGAGATGCTTAC	120
Db	61	TGTGTCTCTGCACAGACCTCATCCAAGATTGTGGTTATGACAGAGGTGGAGATGCTTAC	120
OY	121	TGCACAGCCTGCGCTTCCTGCGAGGTACAAGAACAGCAGCTGGGGGCCACCACAAATGTCAGAGT	180
Db	121	TGCACAGCCTGCGCTTCCTGCGAGGTACAAGAACAGCAGCTGGGGGCCACCACAAATGTCAGAGT	180
OY	181	TGCACAGCCTGCGCTTCCTGCGAGGTACAAGAACAGCAGCTGGGGGCCACCACAAATGTCAGAGT	240
Db	181	TGCACAGCCTGCGCTTCCTGCGAGGTACAAGAACAGCAGCTGGGGGCCACCACAAATGTCAGAGT	240
OY	241	GCTGTCTGTGGGAGCATGTTTGCCAGGTTCATCCGAAGAAGCACGATTTGGAGCGCTGCAG	300
Db	241	GCTGTCTGTGGGAGCATGTTTGCCAGGTTCATCCGAAGAAGCACGATTTGGAGCGCTGCAG	300
OY	301	GACCAAGATGTCATCCCCTGCACGAAGACAGACCCCCACTCTGAGATTCAATGTCCCTTC	360
Db	301	GACCAAGATGTCATCCCCTGCACGAAGACAGACCCCCACTCTGAGATTCAATGTCCCTTC	360
OY	361	CAGTTGACCTTAAGTGGAGGACAGATCCACCACAGTCCCGCCCTCAGAGAGCCACTTGT	420
Db	361	CAGTTGACCTTAAGTGGAGGACAGATCCACCACAGTCCCGCCCTCAGAGAGCCACTTGT	420
OY	421	GCACGTGTGAGCAGCAGCTGCTAGTGG	445
Db	421	GCACGTGTGAGCAGCAGCTGCTAGTGG	445
OY	421	GCACGTGTGAGCAGCAGCTGCTAGTGG	445
Db	421	GCACGTGTGAGCAGCAGCTGCTAGTGG	445
RESULT 9			
LOCUS	AX127574	519 bp	DNA Linear PAT 15-MAY-2001
DEFINITION	Sequence 39 from Patent WO0130850.		
VERSION	AX127574.1 GI:14134266		
KEYWORDS	.		
SOURCE ORGANISM	synthetic construct. artificial sequence.		
REFERENCE	1 (bases 1 to 519)		
AUTHORS	Xu,W., Lofton-Day,C.E., Henne,R., Yao,Y., Novak,J.E., Foster,D.C.		
TITLE	Umlr polyptides		
JOURNAL	Patent: WO 0130850-A 39 03-MAY-2001;		
FEATURES	ZymoGenetics, Inc. (US) Location/Oualifiers 1..519 /organism="synthetic construct" /db_xref="taxon:32630"		
BASE COUNT	misc.feature 1..519 /note="degenerate polynucleotide sequence"		
ORIGIN	1..519 /note="n = A,T,C or G"		
Query Match	Best Local Similarity 45.5%; Score 315.4; DB: 6; Length 519; Matches 260; Conservative 95; Mismatches 72; Indels 0; Gaps 0;		
OY	1	ATGGATGTCCCAAGAAATGAGTACATGGGACCATAAGGGAGCGGTGTGTACTCCTGCCAACGG	60
Db	1	ATGGATGTCCCAAGAAATGAGTACATGGGACCATAAGGGAGCGGTGTGTACTCCTGCCAACGG	60
OY	61	TGTGTCTCTGCACAGACCTCATCCAAGATTGTGGTTATGACAGAGGTGGAGATGCTTAC	120
Db	61	TGTGTCTCTGCACAGACCTCATCCAAGATTGTGGTTATGACAGAGGTGGAGATGCTTAC	120
OY	121	TGCACAGCCTGCGCTTCCTGCGAGGTACAAGAACAGCAGCTGGGGGCCACCACAAATGTCAGAGT	180
Db	121	TGCACAGCCTGCGCTTCCTGCGAGGTACAAGAACAGCAGCTGGGGGCCACCACAAATGTCAGAGT	180
OY	181	TGCACAGCCTGCGCTTCCTGCGAGGTACAAGAACAGCAGCTGGGGGCCACCACAAATGTCAGAGT	240
Db	181	TGCACAGCCTGCGCTTCCTGCGAGGTACAAGAACAGCAGCTGGGGGCCACCACAAATGTCAGAGT	240
OY	241	GCTGTCTGTGGGAGCATGTTTGCCAGGTTCATCCGAAGAAGCACGATTTGGAGCGCTGCAG	300
Db	241	GCTGTCTGTGGGAGCATGTTTGCCAGGTTCATCCGAAGAAGCACGATTTGGAGCGCTGCAG	300
OY	301	GACCAAGATGTCATCCCCTGCACGAAGACAGACCCCCACTCTGAGATTCAATGTCCCTTC	360
Db	301	GACCAAGATGTCATCCCCTGCACGAAGACAGACCCCCACTCTGAGATTCAATGTCCCTTC	360
OY	361	CAGTTGACCTTAAGTGGAGGACAGATCCACCACAGTCCCGCCCTCAGAGAGCCACTTGT	420
Db	361	CAGTTGACCTTAAGTGGAGGACAGATCCACCACAGTCCCGCCCTCAGAGAGCCACTTGT	420
OY	421	GCACGTGTGAGCAGCAGCTGCTAGTGG	445
Db	421	GCACGTGTGAGCAGCAGCTGCTAGTGG	445
OY	421	GCACGTGTGAGCAGCAGCTGCTAGTGG	445
Db	421	GCACGTGTGAGCAGCAGCTGCTAGTGG	445

Db	121	TGYACNGCNTGTCGCCMCMGMNMGNTATYAAARNSNMSTNGGGGCATCATYAARTGYCARMSN	180
Oy	181	TGCATCACCTGTGCTGTGCATCAATCGTGTTCGAAGAGTCAACTGCACAGCTACTTAAT	240
Db	181	TGYATHACNTGTGCGKSTNAHTHAAYMGNGTNCRAARAAGTMAAATGYACMCCNMSNAA	240
Oy	241	GCTGTCGTGGGGAGCTGTTTGCCAGAGGTTCCTCCGAAGAAGACGCGATTGGAGCCTCGAG	300
Db	241	GGNGTNTGGGGATGTGTTTCCNMGNTTYTAMGNAAACMMGNATHGGNGGNTNCAR	300
Oy	301	GACCAGAAGTGCATCCCGTGCACGAAGACAGACCCCCACCTCTGAGGTTCAATGTGCTTAT	360
Db	301	GAYCARGARTRGTATRCMTGTGYNARARACNCCNACMWSMGARTGNCARIGYGNTTY	360
Oy	361	CAGTTGACCTTAGTGGAGGACAGATGCACCCACAGTGCCCTCCAGAGGCCACTTGT	420
Db	361	CARYTNMSNYTGTNGARGCNAGYCNCNCNCGTNCNCNCARGARCACNANYNGTN	420
Oy	421	GCACGTG 427	
Db	421	GCNYTNG 427	
RESULT	10		
AX127568			
LOCUS	AX127568	528 bp	DNA linear PAT 15-MAY-2001
DEFINITION	Sequence 33 from Patent WO0130850.		
ACCESSION	AX127568		
VERSION	AX127568.1 GI:14134264		
KEYWORDS	human.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 528) Xu,W., Lofton-Day,C.E., Henne,R., Yao,Y., Novak,J.E., Foster,D.C. and Yee,D.P. Unlr polypeptides Patent: WO 0130850-A 33 03-MAY-2001;		
JOURNAL	ZymoGenetics, Inc. (US)		
FEATURES	location/Qualifiers		
source	1..528		
BASE COUNT	119 a 158 c 135 g 116 t		
ORIGIN			
Query Match	37.1%; Score 257; DB 6; Length 528;		
Best Local Similarity	97.8%; Pred. No. 7.5e-64;		
Matches 273; Conservative	0; Mismatches 0; Indels 6; Gaps 1.		
Oy	339	CTCTGAGGTTCAATGTGCTTCACAGTTGAGTGTGAGGACAGATGCACCCACAGTGCC	398
Db	1	CTCTGAGGTTCAATGTGCTTCACAGTTGAGTGTGAGGACAGATGCACCCACAGTGCC	60
Oy	399	CCCTCAGAGGCGACACTTGTGTCACGTGTGAGACACCTCTGATGTGTATACCTGCG	458
Db	61	CCCTCAGAGGCGACACTTGTGTCACGTGTGAGACACCTCTGATGTGTATACCTGCG	120
Oy	459	CTTCCGTGGGGCTCTTCTCTCTACTGCAAGAGATTCTTCAACAGACATGTGCAGCGT--	516
Db	121	CTTCCGTGGGGCTCTTCTCTCTACTGCAAGAGATTCTTCAACAGACATGTGCAGCGT	180
Oy	517	---GAGGATTGTGTCACAGTTGAGGCTGATAAACAGACAAGAGAGGATTCCTCTGCC	572
Db	181	TGCAGAGAGGTTGTCTGCAAGTTGAGGCTGATATAAACAGACAAGAGAGGATTCCTCTGCC	240
Oy	573	CGTGCACCCACGAAGAAGACCAAGTGTGATGCCCAAGT 611	
Db	241	CGTGCACCCACGAAGAAGACCAAGTGTGATGCCCAAGT 279	
RESULT	11		

AXI27566	AXI27566	529 bp	DNA	linear	PAT 15-MAY-2001
LOCUS	Sequence 31 from Patent WO0130850.				
DEFINITION	AXI27566				
ACCESSION	AXI27566				
VERSION	AXI27566.1	GI:14134262			
KEYWORDS	.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
TITLE	1 (bases 1 to 529)				
JOURNAL	Xu,W., Lofton-Day,C.E., Henne,R., Yao,Y., Novak,J.E., Foster,D.C.				
FEATURES	and Yee,D.P. UniPr polypeptides Patent: WO 0130850-A 31 03-MAY-2001; ZymoGenetics, Inc. (US) Location/Qualifiers				
source	1..529				
misc-feature	/organism="Homo sapiens" /db_xref="taxon:9606"				
BASE COUNT	116 a 135 c 126 g 107 t 45 others				
ORIGIN					
Query Match	31.8%; Score 220.4; DB 6; Length 529;				
Best Local Similarity	86.0%; Pred.No. 3.5e-53;				
Matches 240; Conservative 16; Mismatches 21; Indels 2; Gaps 2.					
OY	335 GCACGCTGTAGTTCATATGGCTTCCCATCTGAGCTTAGTGAGGCAGATGCCACAG 394				
DB	6 CAAATNCTGAGNTGNATGCGMTTCAGTTNMGASTRKAGTGCAGCAGATGCASCACAG 65				
OY	395 TSCCC-CCTCAGAGAGGACCACTTGTTCACATGTGTGAGCAGCGCTGTGTTACC 453				
DB	66 TGCCCCGCTCAGAGAGGACACTTGTTCGRTMGTTGAGCAGCGCTGTGTTTACC 125				
OY	454 CTGGCTTCTCTGGGCTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 513				
DB	126 CTGGCTTCTCTGGGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 185				
OY	514 C-GTGAGGATTTCGTGACATTTGAGGCTATTAACAGCAGAAGAGGATCTCTTCCC 572				
DB	186 CGNGAGAGTTTGTCTCACAATTGAGGCTATTAACAGCAGAAGAGGATCTCTTTCY 245				
OY	573 CGTGCCACCCAGCAGAGAGACCACTGTCTGATGCCAAGT 611				
DB	246 CGTGCCACCCAGCAGAGAGACCACTGTCTGATGCCAAGT 284				
RESULT 12					
ALJ353136	192505 bp	DNA	linear	PRI 09-MAR-2001	
LOCUS	Human DNA sequence from clone RPL1-133X18 on chromosome X, complete				
DEFINITION	sequence.				
ACCESSION	ALJ353136				
VERSION	ALJ353136.21	GI:13274342			
KEYWORDS	HTG.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
TITLE	1 (bases 1 to 192505)				
JOURNAL	Health,P. Direct Submission Submitted (08-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk				
COMMENT	On Mar 12, 2001 this sequence version replaced gi:13121368. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission				

corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep. This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/ChrX>
RP11-133K18 is from the library RPC1-11.1 constructed by the group
of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>

This sequence is the entire insert of clone RP11-133K18. The true left end of clone RP3-43304 is at 146164 in this sequence.

FEATURES

source

source	1. 192505 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="X" /clone="RP11-13K18" /clone.lib="RPC1-11.1" 2. 590 /note="match: GSS: Em:A0415618" 8. 241 /note="L1M4 repeat: matches 4301. 4547 of consensus" 9. 465 /note="match: GSS: Em:A2516005" 312. 748 /note="L1M2 repeat: matches 5644. 6304 of consensus" 864. 1686 /note="L1PA11 repeat: matches 5308. 6165 of consensus" complement(2058. 2337) /note="match: GSS: Em:A0076345" complement(5857. 6295) /note="match: GSS: Em:A0456667" 6871. 6916 /note="23 copies 2 mer at 100% conserved" 6947. 7069 /note="MER21A repeat: matches 822. 927 of consensus" 7070. 7365 /note="Alus repeat: matches 1. 296 of consensus" 7366. 7498 /note="MER21A repeat: matches 684. 822 of consensus" 7501. 7661 /note="MER21B repeat: matches 363. 523 of consensus" 7520. 7927 /note="match: GSS: Em:A0021529" 7862. 8255 /note="MER21A repeat: matches 3. 406 of consensus" 8256. 10065 /note="L1M1 repeat: matches -1213. 643 of consensus" 10066. 11582 /note="L1M1 repeat: matches 895. 2792 of consensus" 11574. 12563 /note="L1M3 repeat: matches 4976. 5978 of consensus" 12558. 13476 /note="L1PA15 repeat: matches 5171. 6055 of consensus" 13477. 14163 /note="L1M1 repeat: matches 5616. 6299 of consensus" 14164. 18546 /note="L1PA15 repeat: matches 656. 5171 of consensus" 19058. 19370 /note="L1PA13 repeat: matches -651. -339 of consensus"
--------	--

repeat_region	19380..19427	/note="L1PA13 repeat: matches 6109..6156 of consensus"
repeat_region	20158..20170	/note="L2 repeat: matches 992..1002 of consensus"
repeat_region	20171..20593	/note="MSB repeat: matches 1..426 of consensus"
repeat_region	20594..20722	/note="L2 repeat: matches 1002..1147 of consensus"
repeat_region	20803..21452	/note="L2 repeat: matches 1635..2587 of consensus"
repeat_region	21924..22023	/note="MIR repeat: matches 4..106 of consensus"
repeat_region	23030..23050	/note="L1P4 repeat: matches 5372..5391 of consensus"
repeat_region	23051..23600	/note="L1PA13 repeat: matches 5596..6156 of consensus"
repeat_region	23601..24341	/note="L1P4 repeat: matches 4790..5372 of consensus"
misc_feature	complement(24172..24658)	/note="match: GSS: Em: A0617974"
repeat_region	25450..27192	/note="L1PB1 repeat: matches 4389..6155 of consensus"
repeat_region	27175..28098	/note="L1PB1 repeat: matches 3245..4193 of consensus"
repeat_region	28098..28541	/note="L1 repeat: matches 2176..2620 of consensus"
repeat_region	28544..29065	/note="L1PA4 repeat: matches 5626..6146 of consensus"
repeat_region	29066..29291	/note="L1PA4 repeat: matches 5398..5626 of consensus"
repeat_region	29281..29770	/note="L1M2 repeat: matches 1623..2196 of consensus"
repeat_region	29774..29945	/note="AUSQ4X repeat: matches 134..305 of consensus"
repeat_region	29957..32877	/note="L1PB1 repeat: matches 2930..5805 of consensus"
repeat_region	32882..33025	/note="72 copies 2 mer ta 66% conserved"
repeat_region	33027..33390	/note="L1PB1 repeat: matches 5781..6155 of consensus"
repeat_region	33538..35294	/note="L1M3a repeat: matches 501..1804 of consensus"
repeat_region	35271..35571	/note="L1M3a repeat: matches 745..904 of consensus"
repeat_region	35778..36216	/note="L1M3a repeat: matches 213..655 of consensus"
repeat_region	36287..36471	/note="MIR repeat: matches 57..251 of consensus"
repeat_region	36769..37123	/note="L1ME1 repeat: matches 5597..5884 of consensus"
repeat_region	37140..38332	/note="L1MA3 repeat: matches 5105..6299 of consensus"
repeat_region	38339..38440	/note="L1 repeat: matches 2434..2535 of consensus"
repeat_region	38446..38807	/note="MLT2B repeat: matches 1..374 of consensus"
repeat_region	38879..39755	/note="HERV1 repeat: matches 1..916 of consensus"
repeat_region	39796..40402	/note="L1P2 repeat: matches 5050..5656 of consensus"
repeat_region	40403..40850	/note="L1PA4 repeat: matches 5699..6146 of consensus"
repeat_region	40890..45577	/note="HERV1 repeat: matches 905..5578 of consensus"
repeat_region	45578..45888	/note="AUSX repeat: matches 1..311 of consensus"
repeat_region	45889..45965	/note="HERV1 repeat: matches 5578..5654 of consensus"
repeat_region	45966..46316	/note="MLT2B repeat: matches 1..392 of consensus"
repeat_region	46451..49891	/note="L1M1 repeat: matches -1390..2500 of consensus"
repeat_region	49928..50810	

```

repeat_region /note="L1M4 repeat: matches 3343. .4258 of consensus"
50989. .51126 /note="69 copies 2 mer tt 59% conserved"
repeat_region 51129. .51365
/note="L1P4 repeat: matches 5910. .6146 of consensus"
repeat_region 51367. .52126
/note="L1P2 repeat: matches 4727. .5487 of consensus"
repeat_region 52758. .53940
/note="L1 repeat: matches 4150. .5350 of consensus"
repeat_region 53994. .54310
/note="L1M5 repeat: matches 7581. .7917 of consensus"
repeat_region 54357. .54551
/note="L1P3 repeat: matches 5847. .6127 of consensus"
repeat_region 54560. .54862
/note="AluY repeat: matches 1. .303 of consensus"
repeat_region 55116. .57271
/note="L1P8 repeat: matches 3978. .6163 of consensus"
repeat_region 57372. .57301
/note="L1P4 repeat: matches 5338. .5367 of consensus"
repeat_region 57302. .57608
/note="AluX repeat: matches 1. .306 of consensus"
repeat_region 57609. .57770
/note="L1P4 repeat: matches 5367. .5525 of consensus"
repeat_region 57747. .64039
/note="L1P7 repeat: matches 6. .5917 of consensus"
repeat_region 64050. .64157
/note="54 copies 2 mer at 90% conserved"
repeat_region 64162. .65850
/note="L1P2 repeat: matches 4456. .6146 of consensus"
repeat_region 65847. .66663
/note="L1P2 repeat: matches 3623. .4439 of consensus"
repeat_region 66640. .66882
/note="L1P7 repeat: matches 5894. .6134 of consensus"

```

Query Match 26.2%; Score 181.4; DB 9; Length 192505;

Best Local Similarity 99.5%; Pred. No. 1.2e-41;

Matches 182; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 86 AGGATGGGCTATGAGAGGCTGAGATGCTCTGACAGCCGCTCCTCGCAGCT 145

Db 109481 AGGATGGGCTATGAGAGGCTGAGATGCTCTGACAGCCGCTCCTCGCAGCT 109422

QY 146 ACAAGAGCCTGGGGCCACCAAAATGTCAGAGTTGCATCAGCTGTCATCAATC 205

Db 109421 ACAAGAGCCTGGGGCCACCAAGATGTCAGAGTTGCATCAGCTGTCATCAATC 109362

QY 206 GTGTTCAAGAGTCACTGACACGCTACCTGCTGCTGTGCGGAGCTGTTGCCCA 265

Db 109361 GTGTTCAAGAGTCACTGACACGCTACCTGCTGCTGTGCGGAGCTGTTGCCCA 109302

QY 266 GGT 268

Db 109301 GGT 109299

RESULT 13

AC023560/C 205736 bp DNA linear HTG 12-MAR-2000

LOCUS Homo sapiens chromosome 17 clone RP11-85120 map 17, WORKING DRAFT

DEFINITION AC023560

SEQUENCE, 34 unordered pieces.

ACCESSION AC023560.2 GI:7229913

VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.

KEYWORDS human.

SOURCE ORGANISM

Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 205736)

AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.

TITLE Homo sapiens chromosome 17, clone RP11-85120

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 205736)

AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,

TITLE JOURNAL COMMENT

Anderson,S., Baldwin,J., Barna,N., Beda,F., Boguslavsky,L.,
Boukhalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
Choeppel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeRubeis,K., Dewar,K., Dodge,S., Domino,M., Doyle,M.,
Fenster,J., Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karlas,A.,
Klein,J., Landers,T., Largocque,K., Lehoczy,J., Levine,R.,
Lieu,C., Liu,G., Locke,K., MacDonald,P., Margis,N., McCarthy,M.,
McMan,P., McGurk,A., McKernan,K., McPheters,R., Meldrum,D.,
Meneus,L., Mihova,T., Miranda,C., Mienna,V., Morrow,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,T.M.,
Peterson,K., Pierre,N., Pisanl,C., Pollara,V., Raymond,C.,
Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Subramanian,A., Talamas,J., Testaive,S., Theodore,J., Tirell,A.,
Travers,M., Triggillo,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B.,
Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and
Zody,M.

Direct Submission

Submitted (15-FEB-2000) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Mar 12, 2000 this sequence version replaced gi:6978256.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www.seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: 16741

Center clone name: 85_1_20

Sequencing Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 178666 bases at least Q40

Consensus quality: 190940 bases at least Q30

Consensus quality: 197088 bases at least Q20

Insert size: 194000; agarose-fp

Insert size: 202436; sum-of-ctnigs

Quality coverage: 4.2 in Q20 bases; agarose-fp

Quality coverage: 4.0 in Q20 bases; sum-of-ctnigs

NOTE: This is a 'working draft' sequence. It currently

consists of 34 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 1267: contig of 1267 bp in length

1268 1367: gap of 100 bp

1368 2402: contig of 1035 bp in length

2403 2502: gap of 100 bp

2503 3524: contig of 1022 bp in length

3525 3624: gap of 100 bp

3625 4913: contig of 1289 bp in length

4914 5013: gap of 100 bp

5014 6289: contig of 1276 bp in length

6290 6389: gap of 100 bp

6390 7655: contig of 1266 bp in length

7656 7755: gap of 100 bp

7756 9267: contig of 1512 bp in length

9268 9367: gap of 100 bp

9368 11138: contig of 1771 bp in length

11139 11238: gap of 100 bp

11239 13822: contig of 2584 bp in length

13823 13922: gap of 100 bp

```
* 13923 16968: contig of 3046 bp in length
* 16969 17068: gap of 100 bp
* 17069 20185: contig of 3117 bp in length
* 20186 20285: gap of 100 bp
* 20286 22634: contig of 2349 bp in length
* 22635 22734: gap of 100 bp
* 22735 27012: contig of 4278 bp in length
* 27013 27112: gap of 100 bp
* 27113 29342: contig of 2230 bp in length
* 29343 29442: gap of 100 bp
* 29443 33381: contig of 3939 bp in length
* 33382 33481: gap of 100 bp
* 33482 36562: contig of 3081 bp in length
* 36563 36662: gap of 100 bp
* 36663 40982: contig of 4320 bp in length
* 40983 41083: gap of 100 bp
* 41083 45520: contig of 4438 bp in length
* 45521 45620: gap of 100 bp
* 45621 50623: contig of 5003 bp in length
* 50624 50723: gap of 100 bp
* 50724 55610: contig of 4887 bp in length
* 55611 55710: gap of 100 bp
* 55711 61038: contig of 5328 bp in length
* 61039 61138: gap of 100 bp
* 61139 66159: contig of 5021 bp in length
* 66160 66259: gap of 100 bp
* 66260 73479: contig of 7220 bp in length
* 73480 73579: gap of 100 bp
* 73580 80554: contig of 6975 bp in length
* 80555 80654: gap of 100 bp
* 80655 89313: contig of 8659 bp in length
* 89314 89413: gap of 100 bp
* 89414 98132: contig of 8719 bp in length
* 98133 98232: gap of 100 bp
* 98233 107469: contig of 9237 bp in length
* 107470 107569: gap of 100 bp
* 107570 117873: contig of 10304 bp in length
* 117874 117973: gap of 100 bp
* 117974 130973: contig of 13000 bp in length
* 130974 131073: gap of 100 bp
* 131074 143236: contig of 12163 bp in length
* 143237 143336: gap of 100 bp
* 143337 154478: contig of 11142 bp in length
* 154479 154578: gap of 100 bp
* 154579 169967: contig of 15389 bp in length
* 169968 170067: gap of 100 bp
* 170068 187347: contig of 17280 bp in length
* 187348 187447: gap of 100 bp
* 187448 205736: contig of 18289 bp in length.
Location/Qualifiers
1. 205736
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-85120"
/clone.lib="RP11 Human Male BAC"
1. 1267
/misc_feature /note="assembly_fragment"
1368. 2402
/misc_feature /note="assembly_fragment"
2503. 3524
/misc_feature /note="assembly_fragment"
3625. 4913
/misc_feature /note="assembly_fragment"
5014. 6289
/misc_feature /note="assembly_fragment"
6390. 7655
/misc_feature /note="assembly_fragment"
7756. 9267
/misc_feature /note="assembly_fragment"
9368. 11138
/misc_feature /note="assembly_fragment"
```

```
misc_feature 11239. 13822
/misc_feature /note="assembly_fragment"
misc_feature 13923. 16968
/misc_feature /note="assembly_fragment"
misc_feature 17069. 20185
/misc_feature /note="assembly_fragment"
misc_feature 20286. 22634
/misc_feature /note="assembly_fragment"
misc_feature 22735. 27012
/misc_feature /note="assembly_fragment"
misc_feature 27113. 29342
/misc_feature /note="assembly_fragment"
misc_feature 29443. 33381
/misc_feature /note="assembly_fragment"
misc_feature 33482. 36562
/misc_feature /note="assembly_fragment"
misc_feature 36663. 40982
/misc_feature /note="assembly_fragment"
misc_feature 41083. 45520
/misc_feature /note="assembly_fragment"
misc_feature 45621. 50623
/misc_feature /note="assembly_fragment"
misc_feature 50724. 55610
/misc_feature /note="assembly_fragment"
misc_feature 55711. 61038
/misc_feature /note="assembly_fragment"
misc_feature 61139. 66159
/misc_feature /note="assembly_fragment"
misc_feature 66260. 73479
/misc_feature /note="assembly_fragment"
misc_feature 73580. 80554
/misc_feature /note="assembly_fragment"
misc_feature 80655. 89313
/misc_feature /note="assembly_fragment"
misc_feature 89414. 98132
/misc_feature /note="assembly_fragment"
```

Query Match 26.2% Score 181.4; DB 2; Length 205736;
Best Local Similarity 99.5%; Pred. No. 1.2e-41;
Matches 182; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 86 AGGATTGCTTATGAGAGGGGTGAGATGCTTACTGCACAGCCCTCTCTCCAGGT 145
|||||
DB 79395 AGGATTGCTTATGAGAGGGGTGAGATGCTTACTGCACAGCCCTCTCTCCAGGT 79336
|||||
QY 146 ACAAAAGCAGCTGGGGCCACACAAATGTCAGAGTTGCATCAGCTGTGTCATCAATC 205
|||||
DB 79335 ACAAAAGCAGCTGGGGCCACACAGATGTCAAGTTGCATCAGCTGTGTCATCAATC 79276
|||||
QY 206 GTGTTCAAGAGTCACTGCACAGCTTACTTATGCTGTCTGGGAGCTGTTGCCCA 265
|||||
DB 79275 GTGTTCAAGAGTCACTGCACAGCTTACTTATGCTGTCTGGGAGCTGTTGCCCA 79216
|||||
QY 266 GGT 268
DB 79215 GGT 79213
```

RESULT 14
AB040433
LOCUS AB040433 744 bp mRNA linear ROD 22-JUL-2000
DEFINITION Mus musculus mRNA for dtroy, complete cds.
ACCESSION AB040433
VERSION AB040433.1 GI:9392327
KEYWORDS dtroy.
SOURCE Mus musculus CDNA to mRNA.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)
AUTHORS Kojima,T., Morikawa,Y., Copeland,N.G., Gilbert,D.J., Jenkins,N.A.,

THIS PAGE BLANK (USPTO)